THE EVOLUTION OF...

Edited by **Abe Shenitzer**
Mathematics, York University, North York, Ontario M3J 1P3, Canada

Does Mathematics Distinguish Certain Dimensions of Spaces?

Zdzisław Pogoda and Leszek M. Sokołowski

translated by Abe Shenitzer with the editorial assistance of Hardy Grant

PART II

5. THE POINCARÉ CONJECTURE. This conjecture is an important problem in the topology of manifolds and is linked with the classification problem. The great Henri Poincaré created algebraic topology. In particular, he provided in the 1890s the foundations of homotopy and homology theories. To explain the Poincaré conjecture we must say something about the first of these theories.

We assume that the reader is familiar with the definition of the *fundamental group*, or *first homotopy group* $\pi_1(X)$ of a space X. The concept of a fundamental group is a typical example of the procedures used in algebraic topology: given a topological space, we associate with it various groups and reduce the study of the space to the study of these groups. What makes the fundamental group important is that it is a topological invariant, i.e., the fundamental groups of homeomorphic spaces are isomorphic.

In general, the problem of finding the fundamental group of a space is very difficult. Below we give the fundamental groups of some simple spaces.

1. $\mathbb{R}^n$, i.e., n-dimensional Euclidean space, is simply connected. Define an n-dimensional sphere S^n as all the points in $\mathbb{R}^{n+1}$ such that

$$(x^1)^2 + (x^2)^2 + \cdots + (x^{n+1})^2 = 1.$$

For $n \geq 2$, S^n is also simply connected. This means that the fundamental groups of S^n for $n \geq 2$ and of $\mathbb{R}^n$ for all n are trivial.

It would seem that a space is simply connected only if it has no "holes," for then every closed loop can be contracted to a point. This can be translated into the statement that the fundamental group, which counts inequivalent closed loops, counts "holes" as well. This is not quite true. The sphere S^2 minus a point is simply connected (it is homeomorphic to the plane) and so is $\mathbb{R}^3$ minus a point (or even minus a ball). In both cases the fundamental group fails to "notice" the hole.

2. The fundamental group of the circle S^1 is isomorphic to the additive group $\mathbb{Z}$ of the integers. Indeed, if $n \neq m$, traversing the circle n times yields a loop inequivalent to that obtained by traversing it m times. Other spaces

with the same fundamental group are the plane minus a point, and, more generally, any figure in the form of a ring or a cylinder. Yet another space with the same fundamental group is $\mathbb{R}^3$ minus a straight line.

3. On a torus $S^1 \times S^1$ there are three inequivalent types of closed loops: contractible closed loops, closed loops in a plane perpendicular to the axis of the torus, and closed loops in the plane of the axis of the torus. It follows that the fundamental group of the torus is the direct sum $\mathbb{Z} \oplus \mathbb{Z}$. Another space with the same fundamental group is $\mathbb{R}^3$ minus two linked circles.

Back to our main topic. Around 1895 Poincaré first advanced a conjecture that he formulated rigorously in 1904. The conjecture was to the effect that *every simply connected compact 3-manifold without a boundary is homeomorphic to S^3.*

The conjecture sounds deceptively simple. Mathematicians have pondered it for close to a century without being able to prove it or to disprove it. In the process they generalized it to an arbitrary number of dimensions. This is not a trivial generalization, for the conjecture is false for manifolds that are just simply connected (for example: $S^2 \times S^2$ is a compact and simply connected manifold without boundary which is *not* homeomorphic to S^4). This led to the following version of the generalized Poincaré conjecture for topological manifolds: *a compact n-manifold without boundary that has the homotopy type of S^n is homeomorphic to S^n.*

We digress to explain what is meant by the "homotopic sameness" of spaces. (This concept is somewhat similar to that of the homotopic sameness of paths.) The definitions that follow are taken from **[11]**. I denotes $[0, 1]$.

Definition. Two continuous maps $\varphi_0, \varphi_1 : X \to Y$ are *homotopic* if and only if there exists a continuous map $\varphi : X \times I \to Y$ such that, for $x \in X$,

$$\varphi(x,0) = \varphi_0(x),$$

$$\varphi(x,1) = \varphi_1(x).$$

If two maps φ_0 and φ_1 are homotopic, we shall denote this by $\varphi_0 \simeq \varphi_1$. This is an equivalence relation. The equivalence classes are called *homotopy classes* of maps.

To better visualize the geometric content of the definition, let us write $\varphi_t(x) = \varphi(x, t)$ for any $(x, t) \in X \times I$. Then, for any $t \in I$,

$$\varphi_t : X \to Y$$

is a continuous map. Think of the parameter t as representing time. Then, at time $t = 0$, we have the map φ_0, and, as t varies, the map φ_t varies *continuously* so that at time $t = 1$ we have the map φ_1. For this reason a homotopy is often spoken of as a continuous deformation of a map **[11**, p. 64].

We can now define the "homotopic sameness" of spaces.

Definition. Two spaces X and Y are of the *same homotopy type* if there exist continuous maps (called *homotopy equivalences*) f: $X \to Y$, g: $Y \to X$ such that $gf \simeq$ identity: $X \to X$ and $fg \simeq$ identity: $Y \to Y$ **[11**, p. 84].

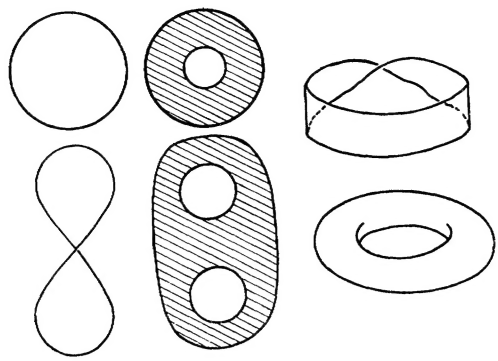

Figure 4. A ring, a Möbius band, and a circle are homotopically equivalent. So too are a figure eight and a "ring" with two holes. A figure eight and a circle are not homotopically equivalent. A sphere and a plane have the same fundamental groups but are not homotopically equivalent. A sphere minus a point and a plane are homotopically equivalent (and even homeomorphic). A torus minus a point, a plane minus two points, and a figure eight are homotopically equivalent.

Before continuing, we recall several definitions of terms that appear in the sequel. (A useful reference is **[27]**.) We assume familiarity with the notion of a topological manifold and of standard terms associated with this notion.

α. A smooth manifold, also called a C^∞-manifold, is a manifold with C^∞ transition functions.

β. A homeomorphism between open subsets of Euclidean space is PL (piecewise linear) if there is a triangulation of each open set into rectilinear simplexes such that the homeomorphism is linear on each simplex.

A manifold is PL if it admits a PL structure. A PL structure on a manifold is specified by an atlas such that the transition functions are all PL-homeomorphisms (such homeomorphisms between manifolds are defined in the usual way by using charts).

γ. If M and N are PL manifolds and f: $M \to N$ and f^{-1} are PL-homeomorphisms, then we say that M and N are PL-isomorphic.

δ. Let M and N be smooth manifolds. f: $M \to N$ is called a C^∞-homeomorphism if, for each $x \in M$, there are charts (U, φ) and (V, ψ) in the atlases of M and N, respectively, such that $x \in U$, $f(U) \subseteq V$, and $\psi \circ f \circ \varphi^{-1}$ is a C^∞-map between open subsets of $\mathbb{R}^n$. f is called a diffeomorphism from M to N if both f and f^{-1} are C^∞-homeomorphisms.

By now we are prepared to resume our account.

The generalized conjecture for topological manifolds was first proved for large values of n **[12]**. Between 1960 and 1962 Stephen Smale proved it for smooth manifolds of dimension $n \geq 5$ **[13]** while John Stallings and Christopher Zeeman proved it for PL manifolds of dimension $n \geq 5$. The only open cases left were $n = 3$ and 4. (For $n = 2$ the generalized conjecture follows immediately from the theorem on the classification of compact surfaces without boundary. The only such simply connected surface is S^2.) In 1964 Zeeman proved the following variant of the generalized conjecture: A compact PL-manifold M^n, $n \geq 6$, without boundary that has the homotopy type of S^n is PL-isomorphic to S^n.

These results were obtained by using powerful new methods for the study of topological spaces discovered in the 1950s. But their use is limited to manifolds of large dimension; there is "not enough room" in low-dimensional manifolds for these methods to work. That is why, surprisingly, the topology of low-dimensional manifolds is more difficult to study than the topology of high-dimensional ones; the techniques used to study the latter are not applicable to the former.

We note that, depending on the structures that can be introduced on manifolds, there are three formulations of the generalized Poincaré conjecture. The one stated earliest applies to topological manifolds. In the version for differentiable manifolds all the mappings are smooth and the homeomorphisms are replaced by diffeomorphisms. In the version for PL manifolds the mappings, homotopies, and isomorphisms are piecewise linear.

Here is what was known about the generalized Poincaré conjecture before 1982:

a. The topological case: the conjecture is known to be true for all dimensions other than the distinguished dimensions 3 and 4.
b. The PL case: the conjecture is known to be true for all dimensions other than the distinguished dimensions 3 and 4.
c. The differentiable case: the conjecture is known to be true for $n = 5$ and 6. In general, it is false for $n \geq 7$. This is a consequence of John Milnor's discovery **[14]** of the existence of "nondiffeomorphic" differentiable structures on spheres of dimension $n \geq 7$ (see Section 6).

In 1981 M. H. Freedman proved the topological variant of the conjecture for the case $n = 4$ **[15]**. In 1986 Colin Rourke and Eduardo Rego announced a proof of the classical Poincaré conjecture for $n = 3$. However, it turned out that their proof contained a technical error that could not be eliminated.

6. DIFFEOMORPHIC AND NONDIFFEOMORPHIC DIFFERENTIABLE STRUCTURES. We recall the definition of a differentiable structure on a smooth manifold M: By an atlas on M we mean a set of charts (U, φ) such that the U's cover M. Two atlases are said to be *compatible* if their union is an atlas. Alternatively, two atlases are compatible if, whenever a chart (U, φ) in one atlas and a chart (V, ψ) in the other atlas satisfy $U \cap V \neq \varnothing$, then the composition

$$\varphi \circ \psi^{-1} : \psi(U \cap V) \to \varphi(U \cap V)$$

is a diffeomorphism. Compatibility of atlases is an equivalence relation, and an equivalence class of atlases is called a *differentiable structure* on M. Briefly, we say that two different atlases define the same differentiable structure if they are compatible in this sense.

It is a complete triviality to note that any differentiable manifold admits many distinct differentiable structures. In fact, let h be a completely arbitrary homeomorphism from the manifold M to itself. If the given differentiable structure is described by the atlas $\{(U, \varphi)\}$, then we can use h to construct a new atlas $\{(U, h \circ \varphi)\}$ and hence a new differentiable structure. This new differentiable structure will coincide with the original one only if h is actually a diffeomorphism.

Two differentiable structures that are related in this trivial way are said to be "diffeomorphic." That is, two differentiable structures on a topological manifold M are diffeomorphic if there is a homeomorphism from M to M that becomes a diffeomorphism when these structures are utilized.

As an example, if our manifold M is the real line $\mathbb{R}$ with its usual differentiable structure, then the homeomorphism

$$h(x) = x^3$$

does not have a differentiable inverse, and hence gives rise to a new differentiable structure. (This h is a diffeomorphism except at the origin, but one can equally well construct examples that are *nowhere* differentiable.)

Until 1956 it was generally assumed that this was the only way to generate distinct differentiable structures. In other words, it was expected that two differentiable manifolds that are homeomorphic would necessarily be diffeomorphic. Small wonder that John Milnor's discovery **[14]** that S^7 admits as many as 28 nondiffeomorphic differentiable structures was viewed as sensational. Soon mathematicians found out that, in this respect, S^7 is not an oddity, for higher-dimensional spheres have similar properties **[16]**. It was shown that it is possible to construct nondiffeomorphic differentiable structures on other manifolds as well. This was a great triumph of the methods of algebraic topology. A property viewed for a long time as natural turned out to be exceptional as soon as one looked at manifolds of sufficiently high dimension. Small wonder that people asked the obvious question: Is it possible to construct on $\mathbb{R}^n$ a structure nondiffeomorphic to the natural one? It required great effort to prove that for $n \neq 4$ the answer is negative. Mathematicians breathed more freely—at least here there were no surprises **[17]**. But what about $\mathbb{R}^4$? People were prepared for technical difficulties but assumed that more powerful methods would show that in this respect $\mathbb{R}^4$ resembles other $\mathbb{R}^n$. Not so! Freedman's papers on the Poincaré conjecture for dimension 4 made it possible for S. K. Donaldson to prove that there are nondiffeomorphic structures on $\mathbb{R}^4$ **[15]**, **[18]**; one such is known as the *exotic* $\mathbb{R}^4$. Moreover, Donaldson and R. Gompf found other nondiffeomorphic structures **[18]**, **[19]**. But many mathematicians were surprised when Gompf showed **[20]** that there is a countably infinite collection of nondiffeomorphic differentiable structures on $\mathbb{R}^4$. Then C. H. Taubes **[28]** proved that there are *uncountably many* such structures. Departures from the norm are one thing, but such extremism on the part of $\mathbb{R}^4$ . . . .

In this connection it is worth noting a fact that in the future may become more than just a curiosity. For most of the 20th century mathematics may be said to have parted company with physics, and situations in which physical problems inspire fundamental mathematical research are few and far between. One well-known exception is the rise of the theory of distributions, which was the mathematicians' response to the introduction by Dirac of the δ function into quantum mechanics. It turns out that the methods used by Donaldson and Gompf are another such exception. They used results and methods from various areas of mathematics, such as algebraic topology, differential geometry, and algebraic geometry, but they also used results and methods from the theory of gauge fields, a theory hitherto of interest to physicists alone. The application of methods of this theory was an essential element of their construction, and was viewed as a sensational departure by mathematicians unaccustomed to seeing purely mathematical problems solved by methods borrowed from theoretical physics. Given the present state of mathematical knowledge, Gompf and Donaldson could not have obtained their results if they were restricted to the "traditional" methods of algebraic topology. The theory of gauge fields bids fair to become the darling of topologists **[18]**, **[21]**, **[22]**, **[23]**.

What is the "meaning" of nondiffeomorphic differentiable structures? The following is a partial answer to this question.

Let $\mathscr{R}^4$ denote $\mathbb{R}^4$ with the "exotic" differentiable structure found by Donaldson. In this differentiable manifold one can find a compact set that cannot be surrounded by any *smoothly* embedded 3-sphere! It is easy to find *continuously* embedded 3-spheres: for example, choose any $\mathbb{R}^4$ metric and look at the 3-sphere $S(r)$ centered at the origin with radius r (i.e., the set of all points x in $\mathscr{R}^4$ with $|x| = r$). However, for r big enough $S(r)$ will be very jagged! This is very different from how our familiar $\mathbb{R}^4$ works, and indeed suffices to show that $\mathscr{R}^4$ has a differentiable structure nondiffeomorphic to that of the usual $\mathbb{R}^4$ **[26]**.

A natural and complementary question to the one just asked is whether it is always possible to put a compatible differentiable structure on a topological manifold; in other words, whether it is always possible to define the notion of a *differentiable* function on a topological manifold in such a way that it is necessarily *continuous*. In dimensions 1, 2, and 3 the answer is always yes; in fact, each manifold of dimension 1, 2, and 3 admits just one differentiable structure (of course, up to a diffeomorphism) compatible with the topological structure. In dimensions 4 and higher the answer is no; for example, there are infinitely many compact 4-manifolds that admit no differentiable structure **[26]**. Similarly, it is known that any differentiable manifold can be given a compatible PL structure (i.e., it can be triangulated), and for dimensions $n \leq 6$, every PL n-manifold admits a unique (up to a diffeomorphism) compatible differentiable structure.

7. AN ATTEMPT AT A SUMMARY. It can be argued that the preceding examples have been tendentiously selected, and that one can find just as many—in fact more—theorems that behave in a "proper" manner in all dimensions. Moreover, for an arbitrary fixed n it is possible to find pathological situations precisely in n-dimensional space. Homotopy theory supplies many relevant examples.

The area of homotopy theory we have in mind is the one concerned with multidimensional knots and links. This theory arose as a natural generalization of 1-dimensional knots and links. An n-dimensional knot is homeomorphic to a sphere S^n, whereas an n-dimensional link is a disjoint sum of a finite number of knots. High-dimensional links can behave very differently. Consider, for example, a link made up of two spheres of dimension 50. In a space of dimension 102 or higher such links are splittable. Spaces of dimension 101, 100, 99, and 98 admit a nontrivial link, but spaces of dimension 97 and 96 don't! Nontrivial links can again be constructed in spaces whose dimension lies between 96 and 52 **[24]**.

Can we nevertheless claim that mathematics singles out dimensions 3 and 4? Our cautious reply, dictated by the imprecision of terms such as "singles out" and "distinguishes" and by the inescapable element of subjectivism present in all interpretations, is that it does.

Unlike the theories of knots and links, the preceding problems belong to the most fundamental problems of topology, and, more generally, of mathematics. It is no accident that the topology of low-dimensional spaces (including the distinguished dimensions) has become an important branch of topology that relies on specific research methods. A surprising variety of difficulties, associated with low-dimensional versions of a great many theorems, forced mathematicians to create appropriate new methods and approaches. Specialization went even further: there came into being a topology of 3-manifolds and a distinct 4-dimensional topology. We emphasize that our discourse is limited to topology, for dimension is a topological concept and not a set-theoretic one.

We give other examples in which the dimensions 3 and 4 are singled out, examples less weighty from a mathematician's standpoint. We know from group

theory that the group of rotations $SO(n)$ of $\mathbb{R}^n$ is simple for all n other than 4, when it is the direct product $SO(3) \times SO(3)$. Next an example from the theory of vector spaces of special importance for physicists: only in $\mathbb{R}^3$ can we define the cross product of vectors, and only there is the curl of a vector a vector. If we lived in a Euclidean (or almost Euclidean) space of another dimension, then the moment of a force and the moment of momentum would be antisymmetric tensors rather than vectors.

The question of how 3-dimensional space is distinguished by physical phenomena is completely outside the scope of our article. Nevertheless it is tempting to mention a few such phenomena. This problem was first investigated by Paul Ehrenfest in a famous article published in 1917 [**25**]. It turns out that the 3-dimensionality of physical space plays a key role for the most important phenomena of nature. If $n \neq 3$, then there are no stable atoms, so that everything connected with chemistry and biology ceases to be; elementary particles can coalesce into larger objects only under the action of gravitational and nuclear forces. But they cannot form planetary systems for the orbits of planets are unstable. Thus such a world would be completely different from the one we know. Only in $\mathbb{R}^3$ can wave phenomena be used for reliable transmission of information. Somewhat later Hermann Weyl showed that only in 4-dimensional spacetime are Maxwell's equations conformally invariant. There are more examples of this kind.

Given the present state of knowledge, there is no connection between problems that incline mathematicians to single out the dimensions 3 and 4 and the fact that physical spacetime—in its entirety or in its macroscopically observable part—is 4-dimensional. All attempts to relate these two facts are outside mathematics and physics, and belong to the realm of free philosophical reflection. The mathematician's assertion that "3- and 4-dimensional spaces are distinguished" is an abbreviated description of the fact that there are many independent and important problems whose nature or solution is atypical in such spaces. In no case can we say that such spaces are distinguished or singular "by their nature." Such a statement goes beyond the confines of mathematics. If anything, the connection with physics is even more baffling. Many physicists think that the deepest foundations of mathematics derive from the physical world. At least some mathematicians doubt this. Even if this were true (how is one to justify such a claim?), we have no idea of how the dimension of the universe we live in is reflected in the mathematical properties we discussed earlier.

On the other hand, the assumption that the coincidence of the dimension of the physical universe and of the dimension of spaces in which many important and independent topological problems have a singular character is entirely accidental and has no deeper foundation is hardly satisfactory. One thing is certain: the question of why the dimension of physical spacetime is what it is, is a genuine scientific question.

ACKNOWLEDGMENTS. This is an abbreviated version, approved by the authors, of a Polish paper titled "Czy matematyka wyróżnia jakiś wymiar przstrzeni?" ("Does mathematics distinguish certain dimensions of spaces?"). The authors are Zdzisław Pogoda and Leszek M. Sokołowski. The paper appeared in *Postępy fizyki* (*Progress of physics*) 40 (1989), 407–433.

The translator wishes to thank Terry Gannon and David Spring for their help.

REFERENCES

1. H. Poincaré, Pourquoi l'espace a trois dimensions?, *Revue de Metaphysique et de Morale* 20 (1912).
2. J. Stillwell, *Classical topology and combinatorial group theory*, second ed., Springer-Verlag, New York, 1993.

3. K. Menger, *Dimensionstheorie*, Teubner, 1928.
4. V. G. Boltyanskiĭ and V. A. Efremovich, *Anschauliche kombinatorische Topologie*, Vieweg, Braunschweig, 1986.
5. R. Engelking, *Dimension theory*, North-Holland, Amsterdam, 1978.
6. H. S. M. Coxeter, *Introduction to geometry*, Wiley, New York, 1969.
7. D. Hilbert and S. Cohn-Vossen, *Geometry and the imagination*, Chelsea, New York, 1952, tr. P. Nemenyi.
8. S. Smale, On the structure of manifolds, *Amer. J. Math.* 84 (1962) 387–399.
9. A. T. Fomenko, *Differential geometry and topology*. Supplementary chapters, Moscow State University, Moscow, 1983. (Russian)
10. A. A. Markov, *Proceedings of the International Congress of Mathematicians*, 1958, 300–306.
11. W. S. Massey, *Algebraic topology: an introduction*, Harcourt, Brace & World, New York, 1967.
12. R. Duda, On the Poincaré conjecture, *Wiadomości Matematyczne* XVIII (1974) 19–39. (Polish)
13. S. Smale, Generalized Poincaré's conjecture in dimensions greater than four, *Ann. Math.* 74 (1961) 391–406.
14. J. Milnor, On manifolds homeomorphic to the 7-sphere, *Ann. of Math.* (2), 64 (1956) 399–405.
15. M. H. Freedman, The topology of four-dimensional manifolds, *J. Differential Geom.* 17 (1982) 357–454.
16. M. A. Kervaire and J. W. Milnor, Groups of homotopy spheres I, *Ann. of Math.* (2) 77 (1963) 504–537.
17. R. Kirby and L. Siebenmann, *Foundational essays on topological manifolds, smoothings and triangulations*, Ann. of Math. Studies 88, Princeton University Press, Princeton, 1977.
18. S. K. Donaldson, An application of gauge theory to four-dimensional topology, *J. Differential Geom.* 18 (1983) 279–315.
19. R. Gompf, Three exotic $\mathscr{R}^4$'s and other anomalies, *J. Differential Geom.* 18 (1983) 317–328.
20. R. Gompf, An infinite set of exotic $\mathscr{R}^4$'s, *J. Differential Geom.* 21 (1985) 283–300.
21. R. J. Stern, Instantons and the topology of 4-manifolds, *Math. Intelligencer* 5 (1983) 39–44.
22. R. J. Stern, Gauge theories as a tool for low-dimensional topologists, 497–507, in: *Perspectives in mathematics*, Birkhäuser, Basel, 1984.
23. H. B. Lawson Jr., *The theory of gauge fields in four dimensions*, CBMS, AMS, Providence, 58 (1985).
24. D. Rolfsen, *Knots and links*, Publish or Perish, Berkeley, 1976.
25. P. Ehrenfest, In what way does it become manifest in fundamental laws of physics that space has three dimensions? *Proc. Amsterdam Acad.* 20 (1917).
26. D. S. Freed, K. K. Uhlenbeck, *Instantons and four-manifolds*, 2nd ed., Springer-Verlag, New York, 1991.
27. *Encyclopedic dictionary of mathematics*, 2nd edition, MIT Press, Cambridge MA, 1987.
28. C. H. Taubes, Gauge Theory on asymptotically periodic 4-manifolds, *J. Differential Geom.* 25 (1987) 363–430.
34. J. H. Schwarz, Superstrings, *Phys. Today* 40 (November 1987) 33–40.

Note: A glossary of physical terms used in this article may be found in pp. 225–228 of *Superstrings. A theory of everything*? Ed. P. C. W. Davies and J. Brown, Cambridge University Press, Cambridge, 1988.